On the Fourier transform for a symmetric group homogeneous space

Ramakrishna Kakarala

School of Computer Engineering
Nanyang Technological University
Singapore 637798
ramakrishna@ntu.edu.sg

Abstract

By using properties of the Young orthogonal representation, this paper derives a simple form for the Fourier transform of permutations acting on the homogeneous space of n-dimensional vectors, and shows that the transform requires $2n-2$ multiplications and the same number of additions.

Key words: Symmetric group, Fourier transform, complexity.

1. Introduction

Let S_n denote the symmetric group on n elements, and S_n^n the subgroup fixing the n-th element. This paper derives a simplification for the Fourier transform of S_n acting on $I_n = \{1, 2, \ldots, n\}$, or equivalently, the coset space S_n/S_n^n. Fourier analysis of permutations on I_n is important for the statistical analysis of ranked data [1], pattern matching, and other applications.

To put the aim of the paper in context, it is useful to consider the ordinary Fourier transform. Let $\mathcal{H}$ be the $n \times n$ unitary matrix with entries $\mathcal{H}_{k,\ell} = (\sqrt{n})^{-1} e^{-j2\pi(k\ell)/n}$. Then $X = \mathcal{H}x$ is the discrete Fourier transform of the vector x. If Δ_d is the translation operator that sends $n \mapsto (n+d) \bmod n$, and Φ_d is the phase shift matrix $\mathrm{diag}[1, e^{j2\pi d/n}, \ldots, e^{j2\pi d(n-1)/n}]$, then

$$\mathcal{H}\Delta_d\, x = \Phi_d X. \tag{1}$$

Similarly, the permutation Fourier transform presented below converts permutations on I_n to group representation "phase" shifts.

Fast Fourier transforms on the groups S_n and their homogeneous spaces have been studied previously. In particular, by applying the method of Clausen [2], Maslen and Rockmore [3, Thm 6.5] give an upper bound for the number of operations (either multiplications or additions) on S_n/S_n^n as $n^3 - n^2$. Maslen [4, Thm 3.5] improves the bound on the same space to show that, at most, $3n(n-1)/2$ operations are necessary. This paper shows that $2n - 2$ operations are sufficient.

2. Background for this paper

We use standard results for permutations [5]. An adjacent transposition is the permutation $\tau_k = (k, k+1)$ that exchanges the k-th and $(k+1)$-th elements but leaves all others unchanged. Every permutation may be written as a product of adjacent transpositions.

The Fourier transform on S_n relies on the group's irreducible unitary representations, with "frequencies" given by arithmetic partitions. Let $\nu = (n_1, \ldots, n_q)$ be a partition of n with $n_i \geq n_{i+1}$ and $n_1 + \ldots + n_q = n$; we write $\nu \vdash n$. For every $\nu \vdash n$ there exists an *irreducible representation*, denoted D_ν. For example, when $\nu = (n)$, we have $D_{(n)}(\sigma) = 1$ for all $\sigma \in S_n$. For other ν, we use the Young orthogonal representation (YOR) to construct the matrices. The Fourier transform of $f : S_n \to \mathbb{C}$ is

$$F(\nu) = \sum_{\sigma \in S_n} f(\sigma) D_\nu(\sigma), \quad \nu \vdash n. \tag{2}$$

For each ν, the coefficient $F(\nu)$ is a $n_\nu \times n_\nu$ matrix. If $f(\sigma) = g(\delta\sigma)$, i.e., f and g are left translates of each other, then, in a manner similar to (1), we obtain that $G(\nu) = D_\nu(\delta)^t F(\nu)$. Of particular interest in this paper is the "fundamental frequency" of the transform given by the partition $\phi = (n-1, 1)$. The $(n-1)^2$ entries of D_ϕ are obtained from the YOR as described in detail below.

It suffices to describe D_ϕ on the adjacent transpositions $\{\tau_k\}$, for those generate S_n. Let $D_\phi(\tau_1)$ be the $(n-1)$-dimensional matrix $\mathrm{diag}[1, 1, \ldots, 1, -1]$. For any m, let $\mathcal{I}_m$ denote the m-dimensional identity matrix, and for $k = 2, \ldots n-1$, let R_k be the 2×2 symmetric matrix

$$R_k = \begin{bmatrix} -\frac{1}{k} & \sqrt{1 - \frac{1}{k^2}} \\ \sqrt{1 - \frac{1}{k^2}} & \frac{1}{k} \end{bmatrix}. \tag{3}$$

Now, for $k = 2, \ldots, n-1$, define $D_\phi(\tau_k)$ to be the symmetric, block-diagonal, matrix

$$D_\phi(\tau_k) = \begin{bmatrix} \mathcal{I}_{n-k-1} & 0 & 0 \\ 0 & R_k & 0 \\ 0 & 0 & \mathcal{I}_{k-2} \end{bmatrix}. \tag{4}$$

It may be verified that the matrices $\{D_\phi(\tau_k)\}$ satisfy the Coexeter relations [5, pg 88], and generate the irreducible YOR for partition $\phi = (n-1, 1)$. Furthermore, note that the decomposition of each $\sigma \in S_n^n$ into $\{\tau_k\}$ excludes τ_{n-1}. Therefore, from (4), it follows that, with $\oplus$ denoting matrix direct sum and $O_{n-2}(\sigma)$ a $(n-2)$-dimensional orthogonal matrix,

$$D_\phi(\sigma) = 1 \oplus O_{n-2}(\sigma), \quad \text{for } \sigma \in S_n^n. \tag{5}$$

3. Fourier analysis on the homogeneous space

Our goal is to simplify (2) for functions defined on I_n. We may extend each f defined on I_n to a corresponding function $\widetilde{f}$ on S_n by $\widetilde{f}(\sigma) = f(\sigma(n))$. Note that $\widetilde{f}$ is constant on left cosets of S_n^n and, therefore, "band-limited".

Proposition 3.1. *Given any complex-valued function f defined on I_n, the Fourier coefficients $\{\widetilde{F}(\nu)\}$ of the function $\tilde{f}$ on S_n defined by $\widetilde{f}(\sigma) = f(\sigma(n))$ are such that $\widetilde{F}(\nu) = 0$ unless $\nu = (n)$ or $\nu = \phi = (n-1, 1)$.*

Proof. Since $\widetilde{f}(\sigma\delta) = \widetilde{f}(\sigma)$ for $\delta \in S_n^n$, we have by (2) that $\widetilde{F}(\nu) = \widetilde{F}(\nu) D_\nu(\delta)^t$. By averaging both sides over S_n^n, we get $\widetilde{F}(\nu) = \widetilde{F}(\nu) Z(\nu)$ where

$$Z(\nu) = \frac{1}{(n-1)!} \sum_{\delta \in S_n^n} D_\nu(\delta)^t. \tag{6}$$

Now, the Branching Rule [5, Thm 2.8.3] shows that for $\nu = \phi = (n-1, 1)$ and $\nu = (n)$, the representation D_ν reduces on the subgroup S_n^n to contain the constant representation, and that no other irreducible representation does so. By orthogonality, those matrix entries that are not constant on S_n^n must sum to zero over the subgroup. Therefore $Z(\nu) = 0$ if ν is not (n) or ϕ. $\square$

If $Z(\phi)_{i,j}$ is the (i,j)-th element, then from (5) we have $Z(\phi)_{1,1} = 1$, and, by orthogonality, $Z(\phi)_{i,j} = 0$ for all other (i, j). Since $\widetilde{F}(\phi) = \widetilde{F}(\phi) Z(\phi)$ we obtain that $\widetilde{F}(\phi)$ is zero except possibly in the leftmost column. Hence, the Fourier transform (2) need only be calculated for the partition (n), and for

the $n-1$ entries in the left most column of D_ϕ. Let $\mathcal{F}$ denote the the linear transformation taking any n-dimensional vector x on I_n to its n Fourier transform coefficients $\widetilde{X}((n))$, and the leftmost column entries $\widetilde{X}(\phi)_{i,1}$ for $i = 1, 2, \ldots, n-1$. We write

$$\widetilde{X} = \mathcal{F}x \tag{7}$$

to express the transform, now viewed a matrix operation. The transform (7) requires at most n^2 multiplications and $n(n-1)$ additions. We show below that, in fact, $2n-2$ operations of each kind are sufficient.

Our result relies on the following $n \times n$ matrix U, whose shape is similar to a "reverse" upper Hessenberg matrix:

$$U = \begin{pmatrix} +1 & +1 & +1 & \cdots & +1 & +1 \\ -1 & -1 & -1 & \cdots & -1 & (n-1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & 2 & 0 & \cdots & 0 \\ -1 & +1 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{8}$$

Define $A = (UU^t)^{-1/2}$, and let

$$\mathcal{T} = AU. \tag{9}$$

It is easily seen that $\mathcal{T}$ is an orthogonal matrix, and that A is diagonal with entries $\{\alpha_k\}_{k=1}^n$, with $\alpha_1 = A_{1,1} = 1/\sqrt{n}$, and for $k > 1$, we have

$$\alpha_k = A_{k,k} = \frac{1}{\sqrt{(n-k+1)(n-k+2)}}. \tag{10}$$

Let x be any complex-valued $n \times 1$ vector, and let $X = \mathcal{T}x$. To each $\sigma \in S_n$, let the $n \times n$ matrix $P(\sigma)$ be the permutation matrix obtained from the identity $\mathcal{I}_n$ with rows permuted by σ, i.e., $P(\sigma)_{i,j} = \mathcal{I}_{\sigma(i),j}$. Note that $\sigma \mapsto P(\sigma)$ is an antihomomorphism: $P(\sigma\delta) = P(\delta)P(\sigma)$. To see that, note that for any α we have $P(\alpha)e_k = e_{\alpha^{-1}(k)}$ where e_k is $[0, \ldots, 0, 1, 0, \ldots, 0]^t$ with 1 in the k-th position. If x, y are $n \times 1$ vectors, and $y = P(\sigma)x$, then $y(i) = x(\sigma(i))$ since $P(\sigma)$ is the permutation operator on column vectors. We now establish the following result, comparable to eq. (1).

Theorem 3.2. *For every $\sigma \in S_n$ and all $n \times 1$ vectors x, we have that*

$$\mathcal{T}P(\sigma)\,x = \left[1 \oplus D_\phi(\sigma)^t\right] X$$

Proof. We start by proving for any adjacent transposition τ_k that

$$\mathcal{T}P(\tau_k)\mathcal{T}^t = 1 \oplus D_\phi(\tau_k) = 1 \oplus D_\phi(\tau_k)^t \tag{11}$$

Note from (9), (10), the m-th row of $\mathcal{T}$ for $m > 1$ sums to zero, with the form

$$[-\alpha_m, -\alpha_m, \ldots, -\alpha_m, (n-(m-1))\alpha_m, 0, \ldots, 0]. \tag{12}$$

The product $\mathcal{T}P(\tau_k)$ is the same as $\mathcal{T}$ but with columns k, $k+1$ swapped. By (12), we see that the only rows of $\mathcal{T}P(\tau_k)$ that are affected by the column swap are as follows: for $k = 1$, row n is modified; and for $k > 1$, rows $n-(k-1)$, $n-(k-2)$ are modified. Therefore the product $\mathcal{T}P(\tau_k)\mathcal{T}^t$ is the same as the identity $\mathcal{I}$ in all entries with the following exceptions: when $k = 1$, we have that $[\mathcal{T}P(\tau_1)\mathcal{T}^t]_{nn} = -2\alpha_n^2 = -1$; and when $k > 1$, we have that the 2×2 submatrix, whose upper-left corner indices are $(n-(k-1), n-(k-1))$, has the symmetric form

$$\begin{bmatrix} -(k+1)\alpha_{n-(k-1)}^2 & (k^2-1)\alpha_{n-(k-1)}\alpha_{n-(k-2)} \\ (k^2-1)\alpha_{n-(k-1)}\alpha_{n-(k-2)} & (k-1)\alpha_{n-(k-2)}^2 \end{bmatrix} \tag{13}$$

Subsituting from (10), we find that the above simplifies to R_k as defined earlier in (3), thus verifying (11) for $k = 1, 2, \ldots, n-1$.

For the general case, note that every $\sigma \in S_n$ may be written as a product of adjacent transpositions $\sigma = \tau_{k_1} \cdots \tau_{k_m}$. Since $\sigma \mapsto P(\sigma)$ is an anti-homomorphism, we have that

$$P(\sigma) = P(\tau_{k_1} \cdots \tau_{k_m}) = P(\tau_{k_m}) \cdots P(\tau_{k_1}). \tag{14}$$

Applying a similarity transformation with $\mathcal{T}$ yields

$$\mathcal{T}P(\sigma)\mathcal{T}^t = \mathcal{T}P(\tau_{k_m})\mathcal{T}^t \cdots \mathcal{T}P(\tau_{k_1})\mathcal{T}^t. \tag{15}$$

On applying (11) we establish the theorem:

$$\mathcal{T}P(\sigma)\mathcal{T}^t = \left[1 \oplus D_\phi(\tau_{k_m})^t\right] \cdots \left[1 \oplus D_\phi(\tau_{k_1})^t\right] = 1 \oplus D_\phi(\sigma)^t. \tag{16}$$

$\square$

Note that for the Fourier transform in (7), we also have

$$\mathcal{F}P(\sigma)x = [1 \oplus D_\phi(\sigma)]^t \widetilde{X}$$

from the translation property. Since this is true for all vectors x, we must have $\mathcal{F} = [\lambda_1 \mathcal{I}_1 \oplus \lambda_2 \mathcal{I}_{n-2}]\,\mathcal{T}$. To see that, note that $\mathcal{F} = \mathcal{C}\mathcal{T}$ for some matrix $\mathcal{C}$, and, by applying the Theorem above, we see that $\mathcal{C}$ commutes with all matrices $1 \oplus D_\phi$; the result now follows from Schur's lemma [5, pg 23].

3.1. Computation of the transform

The equality $\mathcal{T} = AU$, combined with the matrix structure in (8), simplifies computation. Let $a_x(n) = x(1)$, $a_x(n-1) = x(1) + x(2)$, ..., $a_x(1) = x(1)+x(2)+\cdots+x(n)$. Computing all $\{a_x(k)\}$ values requires $n-1$ additions due to recursion. If $\hat{X} = Ux$ then $\hat{X}(1) = a_x(1)$, $\hat{X}(2) = (n-1)*x(n)-a_x(2)$, ..., $\hat{X}(n) = x(2) - a_x(n)$. Hence, if a_x has been computed, computing $\hat{X}$ requires $(n-2)$ multiplies and $(n-1)$ additions. Now, since $X = \mathcal{T}x = A\hat{X}$, and A is diagonal, we see that computing X from $\hat{X}$ requires an additional n multiplications. In total, computing the transform $X = \mathcal{T}x$ requires $2n-2$ multiplications and $2n-2$ additions. Note that computing $\widetilde{X} = \mathcal{F}x = \mathcal{C}AUx$ does not require any extra computation as we may premultiply the diagonal matrix $\mathcal{C}$ with A.

4. Conclusions

This paper describes a simplification of the Fourier transform on S_n/S_n^n, and shows that the transform requires $2n-2$ multiplications and the same number of additions.

References

[1] P. Diaconis, A gengeneralization of spectral analysis with applications to ranked data, Annals of Statistics 17 (1989) 949–979.

[2] M. Clausen, Fast generalized fourier transforms, Theor. Comput. Sci. 67 (1) (1989) 55–63.

[3] D. K. Maslen, D. N. Rockmore, Generalized ffts - a survey of some recent results, in: DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Groups and Computation, II, L. Finkelstein and W. Kantor (eds.), 1995, pp. 183–237.

[4] D. K. Maslen, The efficient computation of fourier transforms on the symmetric group, Mathematics of Computation 67 (223) (1998) 1121–1147.

[5] B. E. Sagan, The symmetric group: representations, combinatorial algorithms, and symmetric functions, 2nd Edition, Springer, 1946.